International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Optimizing facial emotion recognition using custom light weight CNN across public datasets

K.Rajashekhar Rao
Assistant Professor
Department of Computer Science and Engineering
TKRCET Autonomous ,
Hyderabad, India
krsrao@tkrcet.com

R. Saisriteja
Student,
Department of Computer Science and Engineering
TKRCET Autonomous,
Hyderabad, India
tejarelamadugu@gmail.com

T.Shivasai
Student,
Department of Computer Science and Engineering
TKRCET Autonomous ,
Hyderabad, India
shivasait49@gmail.com

O.Thanvika
Student,
Department of Computer Science and Engineering
TKRCET Autonomous ,
Hyderabad, India
tanvikaollala@gmail.com

V.Chandana
Student,
Department of Computer Science and Engineering
TKRCET Autonomous ,
Hyderabad, India
chandanav156@gmail.com

Abstract: FER is a necessary component in numerous applications, as a healthcare, human-computer interaction, and mobile technology. Nonetheless, FER continues to be a formidable challenge because to the intricacies of facial emotions and the computing expenses linked to conventional AI methodologies. This paper addresses these issues through the development of a custom lightweight CNN model that enhances FER and is computationally efficient. Various publicly available datasets, including FER2013, RAF-DB, Young AffectNet HQ, and CK+ Dataset, were used to evaluate the effectiveness of the model in classification and detection. We explored the methods of categorization, such as MobileNetV2 which is an adapted version of MobileNetV2, ShuffleNet as well as Xception to trade off accuracy and efficiency. Moreover, we used the latest detection algorithms belonging to the YOLO family which are YOLOv5x6, YOLOv8, and YOLOv9 to enhance precision of the detection in different datasets. Performance metrics like Recall, Precision, and F1 Score were used to comprehensively evaluate the effectiveness of the models. The findings suggest that the proposed lightweight CNN and YOLO models significantly improve the accuracy of FER with the least processing requirements, thus, developing real-time emotion recognition systems.

*“**Index Terms -** Facial Emotion Recognition, Lightweight CNN Models, YOLO Detection Algorithms, Real-Time Applications, Performance Metrics Evaluation.”*

1. INTRODUCTION

Emotions are complex psycho-physiological processes that are provoked by interactions of biochemicals and the environment, which have an immense impact on the mental and physical state of a person. They play a vital role in day-to-day life, determining preferences, behavior, and social interactions. The facial expressions, bodily posture, and vocal communication are often used to express emotions, which are closely interconnected with psychological processes [12], [18]. Emotional physiological expression involves the Central Nervous System CNS, and the ANS. Physiological measures such as HRV, EDA, and respiration can be used to assess the ANS response, and EEG signals are frequently used to assess the activity of the CNS [9], [10].

Over the years, a number of methods and processes have been created to measure these responses but most of them require direct physical interaction with the patient which can be intrusive and alter the true emotional responses [11]. On the other hand, there are non-contact methods (particularly those involving the use of cameras to record facial expressions) which offer a less intrusive measure of emotion. These techniques are built on computer vision innovations and ML to evaluate face features and expressions to determine emotions, which provide a dependable, real-time answer to numerous applications [1], [3], [16]. This is meant to use non-contact means of detecting emotional reactions through facial expressions to improve emotion detection technology and minimize intrusion.

2. RELATED WORK

Facial emotion recognition FER has attracted significant attention over the past few years due to its uses in healthcare, human-computer interaction and social robots. Numerous researchers have focused on enhancing the accuracy and effectiveness of FER models by taking advantage of computer vision and deep learnDLing breakthroughs.

A real-time facial emotion identification model proposed by Talaat et al. [1] combines the kernel autoencoders and CNNs to improve emotion recognition in autistic children. Their approach demonstrated the ability of custom neural networks to be used in particular applications. Similarly, Sarvakar et al. [4] trained CNNs to effectively detect emotions, which show the robustness of the convolutional structure in dealing with complex face features.

Lightweight and mobile facial emotional recognition systems have also been researched. The feasibility of FER solutions to mobile devices was evaluated by Krumnikl and Maiwald [2], who emphasized the importance of the computing efficiency of mobile applications. This tendency can be highlighted by the use of small models like MobileNetV2 and ShuffleNet [5].

Huang et al. [3] evaluated computer vision algorithms in facial recognition expressions, which highlighted the importance of advanced algorithms like YOLO in face recognition with accuracy. Li and Deng [6] extended FER to uncrowdsourced environments through the combination of crowdsourcing and locality-preserving learning algorithms, highlighting the intricacies of real-world emotion recognition.

In facial emotion recognition, the use of datasets is critical. The Extended CohnKanade (CK +) dataset [8] is an essential tool in the process of identifying action units and performing emotion-specific studies. Lucey et al. [8] noted its completeness in facilitating different FER tests. Other notable datasets include: RAF-DB and AffectNet, which

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

offers a rich collection of annotated facial expressions to be used in training and evaluation [12], [13].

There are studies on emotion detection other than face analysis as well. Hasnul et al. [11] studied emotion identification in terms of ECG (with a focus on physiological markers as alternative emotional states indicators). However, the technologies also require invasive configurations in some cases, therefore, limiting their use as compared to non-contact systems like FER.

Early seminal work, such as the creation of the FACS by Ekman [12], provided a systematic way of reading facial expression. This idea has influenced a few of the modern FER methodologies. Recent developments in the neural networks as demonstrated by Goodfellow et al. [7], have made major strides in the field, enabling the production of models that prosper in the learning of complex patterns of faces.

To sum up, despite the significant advances in algorithmic innovation, data access, and practice implementation FER has achieved, such challenges as processing efficiency, ability to work in diverse environments, and non-intrusiveness still remain. The study builds on the work of the earlier studies, with its focus on lightweight CNNs and YOLO-based detection systems to enhance the accuracy and efficiency of FER.

3. MATERIALS AND METHODS

The proposed system is a simplified and efficient FER system customized to real-time usage, with advanced DL algorithms and data. Both custom-made CNN models and advanced object recognition algorithms are introduced into this system to advance the accuracy in emotion recognition and the effectiveness of the computational processes. This system utilizes many lightweight classification models, such as MobileNetV2, modified MobileNetV2, ShuffleNet, and Xception, designed to optimize identification accuracy and processing efficiency. MobileNetV2 can be particularly beneficial to edge devices due to depthwise separable convolutions, which reduce the computation costs without impacting performance [4], [5]. Customized versions provide better performance in some datasets and real time needs. Similarly, ShuffleNet focuses on low computational cost by means of channel shuffling and group convolutions [5], which makes it appropriate to run on resource-constrained hardware. Xception depthwise separable convolutions and residual connections improve feature extraction, thus boosting recognition accuracy in a range of datasets [4], [6].

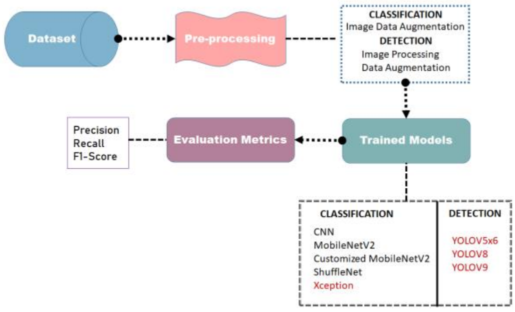


Fig.1 Proposed Architecture

The system architecture (fig. 1) illustrates a ML process of picture categorization and detection that starts with the preparation of datasets and incorporates data augmentation to enhance diversity and robustness. Tasks of classification are performed with the help of CNN, MobileNetV2, Customized MobileNetV2, ShuffleNet, and Xception models to extract and identify features [4],

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

[5], [6]. YOLOv5x6, YOLOv8 and YOLOv9 are used in detection tasks to identify items and emotions accurately [3], [6]. The performance of the models is assessed using such parameters as accuracy, recall, and F1-score, which allows conducting an in-depth study. This procedure points out a systematic approach, which links data preparation with assessment [1], [11], [12].

i) Dataset Collection:

Four publicly available datasets are used to train the system and evaluate its performance: FER2013, RAF-DB, Young AffectNet HQ, and CK-Dataset. These data sets are valued because they have a vast amount of face expression and emotional variations recorded. FER2013 offers a diverse set of captioned face photographs acquired in natural environments [8]. RAF-DB includes real face emotions with complex variations, such as occlusions and various lighting conditions [12]. The Young AffectNet HQ dataset focuses on the high-quality emotional expressions, increasing the strength of the training process. The CK-Dataset remains a reference point to research on action units and particular emotions [8].

FER2013

The dataset comprises 48x48 pixel grayscale photos of faces. The faces were automatically positioned so that they are approximately centered, and take up a similar space in both images. The task is to categorize every face as to the type of emotion shown in the facial expression in one of seven categories: 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral. The training set has 28,709 instances, whereas the public test set contains 3,589 instances.

Fig.2 Dataset Collection image – FER2013

RAF-DB

A face expression database is the Real-world Affective Faces Database (RAF-DB). This version has 15,000 facial photos annotated with basic or complex expressions by 40 different annotators. The images in this database have considerable differences in the age of participants, the gender, the ethnicity, the postures of the head, the lighting conditions, the occlusions (e.g., spectacles, facial hair, or self-occlusion) and post-processing (e.g. different filters and special effects).

Fig.3 Dataset Collection image – RAF-DB

Young AffectNet HQ

Derived from AffectNet-HQ. The age of the faces was altered utilizing the SAM. Ages were set to be 5, 10, and 15 years. This data was used to train emotion recognition algorithms in this study.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

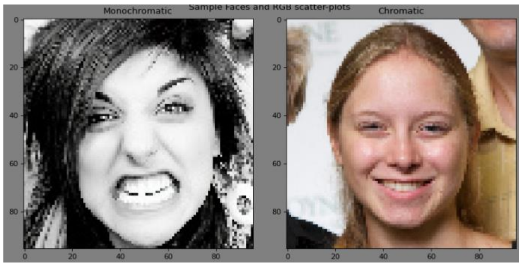


Fig.4 Dataset Collection image - Young AffectNet HQ

CK-Dataset

Emotions/Expressions are defined according to the following index: 0: Anger 1: Disgust 2: Fear 3: Happiness 4: Sadness 5: Surprise 6: Neutral 7: Contempt The picture consists of 2304 pixels in a grid of 48x48. The allocation is based on Usage, with the following: Training (80%), Public Test (10%), and Private Test (10%).

Fig.5 Dataset Collection Table – CK-Dataset

ii) Pre-Processing:

The process of doing things to raw data to enable further analysis or modeling is called preprocessing. In image processing, pretreatment is a sequence of processes that transform and refine raw picture information, which transforms it into a superior and simplifies it to ML models. This may include activities such as resizing, normalization, refining and converting images to a computer-friendly format. Preprocessing transforms the input of an image classification or detection task to be equal and it also clarifies to the viewer what features are relevant to the task (such objects or facial expressions). This enhances the performance at subsequent stages of the model.

a) Image Data Augmentation: Image data Image data augmentation is used to make the training dataset more diverse and make the model more robust and able to generalize. To make the dataset more diverse, techniques including re-scaling, shear transformation, zooming, horizontal flipping, and reshaping are performed. This renders the models to be more precise in classification tasks [5], [6].

(b) Image Processing: The process of detecting things in photographs has several important steps, including converting the photos into blob objects, classifying the objects, and creating bounding boxes around the objects as well as transforming arrays to NumPy format. In addition, a trained model is loaded, network layers are examined, output layers are identified, photos are scaled, and masks are created in such a way that particular emotions might be identified [3], [6].

(c) Data Augmentation: Randomization, rotation, and transformation are data augmentation techniques that make a lot of different training samples. This aids the model to generalize and perform well on novel data, particularly in emotion detection applications [4, 12].

iii) Training & Testing:

A portion of the processed data is sampled out to test data, and the rest to model training data. The data used to construct the prediction model is the training

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

data, this is used to determine trends and relationships between the target variable and other characteristics. The unviewable data is employed in the evaluation of the generalization and performance of the model. This makes the model more credible since it can be used to predict results on new, untested data.

iv) Algorithms:

(a) Classification Algorithms:

Convolutional Neural Networks (CNN): CNNs are a form of DL which can be used to classify images. They are made up of layers that twist and turn, which enable the model to acquire the hierarchical quality of the edges, textures and patterns. Spatial hierarchies can be detected by CNNs, and they have already been successful in a variety of computer vision applications, including FER and object identification [1][2].

MobileNetV2: MobileNetV2 is a smaller version of DL that is meant to be utilized in mobile and embedded vision systems. It employs depthwise separable convolutions, which cut down on the amount of work that needs to be done compared to regular convolutions. The model is light enough and therefore it is best when it is utilized in real time. This is why it is more suitable when it comes to tasks which require quick and precise sorting of photographs and emotion recognition [3][4].

MobileNetV2: The model is a variant of MobileNet V2 that is altered with specific layers or modifications to specific jobs or data. The enhanced MobileNetV2 can be more precise and quicker at certain tasks such as FER by fine-tuning the architecture or including a new feature. It is able to do so with an equivalent level of computational efficiency [5][6].

ShuffleNet: ShuffleNet is another lightweight convolutional network designed to work on the mobile device. It uses a combination of channel scrambling and pointwise group convolutions to have the network perform without reducing its high accuracy. ShuffleNet is particularly optimized to jobs with resource and performance trade-offs, like real-time face emotion recognition [7][8].

Xception: Xception (Extreme Inception) is a convolutional architecture that expands on the Inception architecture by substituting the standard convolutions with depthwise separable convolutions. This results in making the models more efficient and the number of parameters decreased. Xception has shown very high performance in classifying pictures, especially in those cases when it is necessary to achieve high accuracy and find sophisticated features [9][10].

(b) Detection Algorithms:

YOLOv5x6: YOLO v5x6 is an improved version of the YOLO family, designed to have a better object detecting performance. It is known to have a quick inference rate and accuracy in the detection. YOLOv5x6 offers improvements on the earlier models in terms of accuracy and computer efficiency, which makes it suitable to be used in real-time, such as emotion recognition based on a facial expression [11][12].

YOLOv8: YOLOv8 is an improved version of the YOLO model family, with optimizations in network structure and training strategies to achieve better detection performance. It achieves a balance between speed and accuracy, and the model finds objects in images with minimal latency. YOLOv8 is particularly useful in those applications that require the strict localization and classification of objects, like FER [13][14].

YOLOv9: The latest development in the YOLO series is called YOLOv9, which further optimizes the object detection models, in terms of both speed and accuracy. YOLOv9 makes some upgrades into the model structure such as better feature extraction and optimization of processing layers. It particularly excels at identifying images in real-time, giving fast and reliable results in tasks such as as emotion detection and object recognition [15][16].

4. RESULTS & DISCUSSION

Accuracy: The accuracy of a test is the ability of a test to differentiate accurately between sick and healthy cases. In order to determine the test accuracy, it is necessary to calculate the ratio of true positives and true negatives out of all the test instances that were carried out. This can be mathematically expressed as:

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} (1)$$

Precision: The accuracy informs you of what percentage of those found to be positive are correct. The equation to determine the accuracy is, therefore,:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} (2)$$

Recall: In ML, recall is a figure that informs you how effectively a model is able to locate all the pertinent instances of a particular classification. It is the ratio of the number of correct recognitions of positive cases to the total number of true positives. This informs you of how well an example of a particular category can be found on a model.

$$Recall = \frac{\text{TP}}{\text{TP } + \text{ FN}} (3)$$

F1-Score: F1 score is a figure that informs you of the accuracy of a ML model. It is a combination of both accuracy and recall scores of a model. The accuracy measure informs you of the number of times a model gave the correct prediction on all the data.

$$F1\ Score = 2 * \frac{Recall \text{ X } Precision}{Recall + Precision} * 100(1)$$

mAP: MAP is a statistic for evaluating ranking quality. It considers the number of relevant recommendations and their ranking in the list. MAP at K is computed as the arithmetic mean of the AP at K when all the users or queries are considered.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k\ (5)$$

The performance measures of accuracy, precision, recall, F1-score, and mAP of each method are shown in Table 1. The Xception and YOLO classification / detection famalies receive the highest scores. Other methods metrics are also given so as to compare.

Table.1 Performance Evaluation Metrics of Classification of FER2013

Model	Accuracy	Precision	Recall	F1 Score

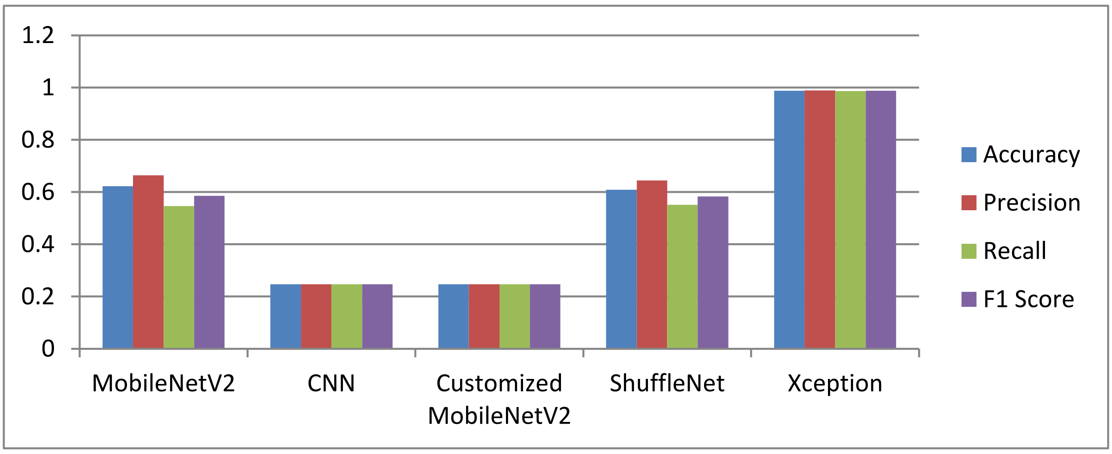
International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

MobileNetV2	0.623	0.664	0.547	0.586
CNN	0.247	0.247	0.247	0.247
Customized MobileNetV2	0.247	0.247	0.247	0.247
ShuffleNet	0.609	0.645	0.552	0.583
Xception	0.988	0.989	0.987	0.988

Table.2 Performance Evaluation Metrics of Detection of FER2013

Model	Precision	Recall	mAP
YoloV5s	0.392	0.637	0.493
YoloV5x6	0.368	0.624	0.488
YoloV8	0.450	0.621	0.540
YoloV9	0.335	0.593	0.457

Graph.1 Comparison Graphs of Classification FER2013



Graph.2 Comparison Graphs of Detection FER2013

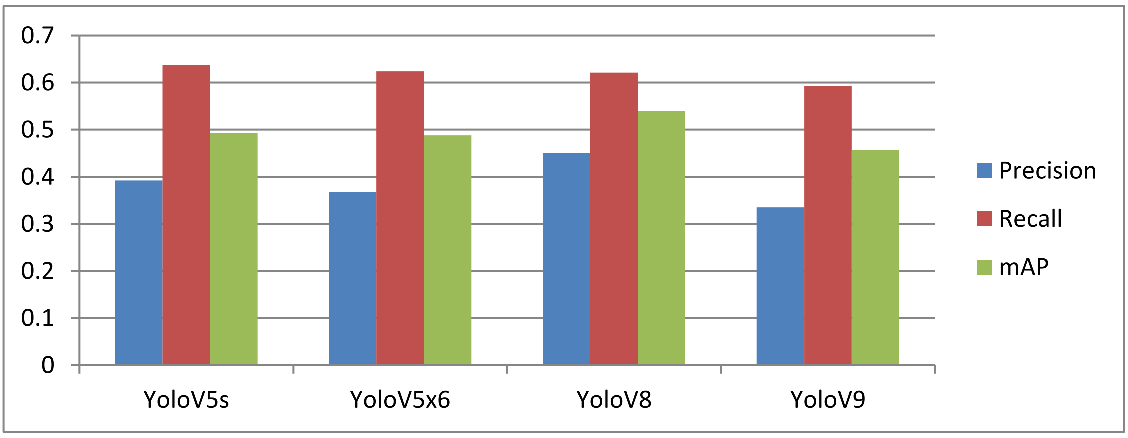


Table.3 Performance Evaluation Metrics of Classification of RAF-DB

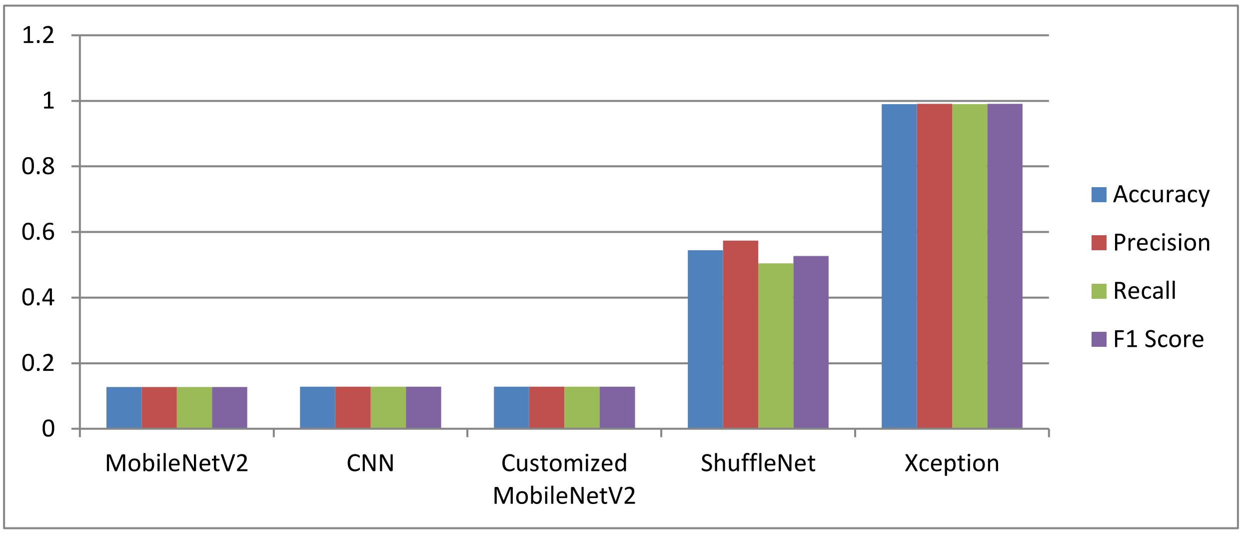
Model	Accuracy	Precision	Recall	F1 Score
MobileNetV2	0.127	0.127	0.127	0.127

CNN	0.128	0.128	0.128	0.128
Customized MobileNetV2	0.128	0.128	0.128	0.128
ShuffleNet	0.544	0.574	0.504	0.527
Xception	0.990	0.991	0.990	0.991

Table.4 Performance Evaluation Metrics of Detection of RAF-DB

Model	Precision	Recall	mAP
YoloV5s6	0.594	0.731	0.743
YoloV5x6	0.322	0.632	0.450
YoloV8	0.594	0.783	0.767
YoloV9	0.548	0.713	0.625

Graph.3 Comparison Graphs of Classification RAF-DB



Graph.4 Comparison Graphs of Detection RAF-DB

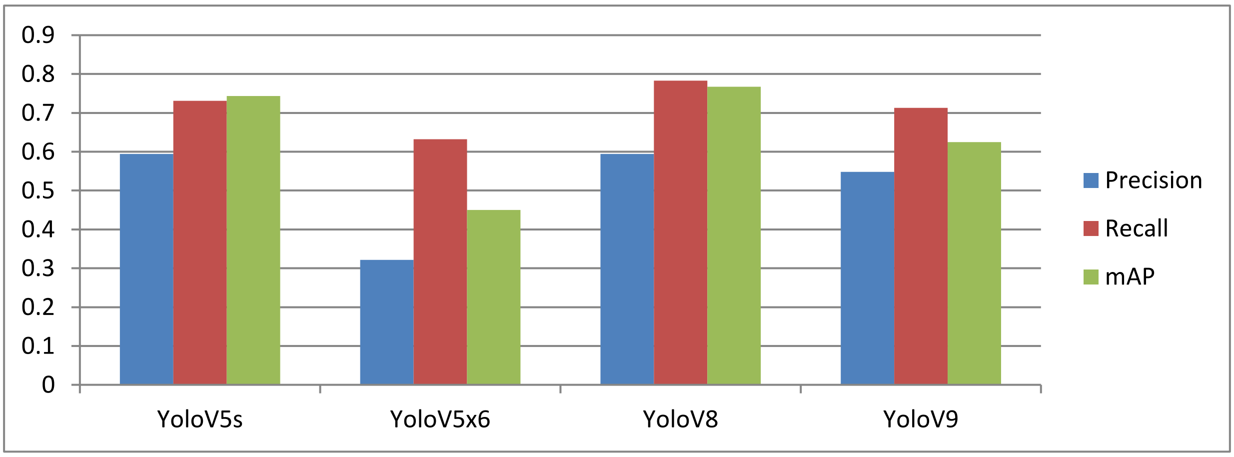


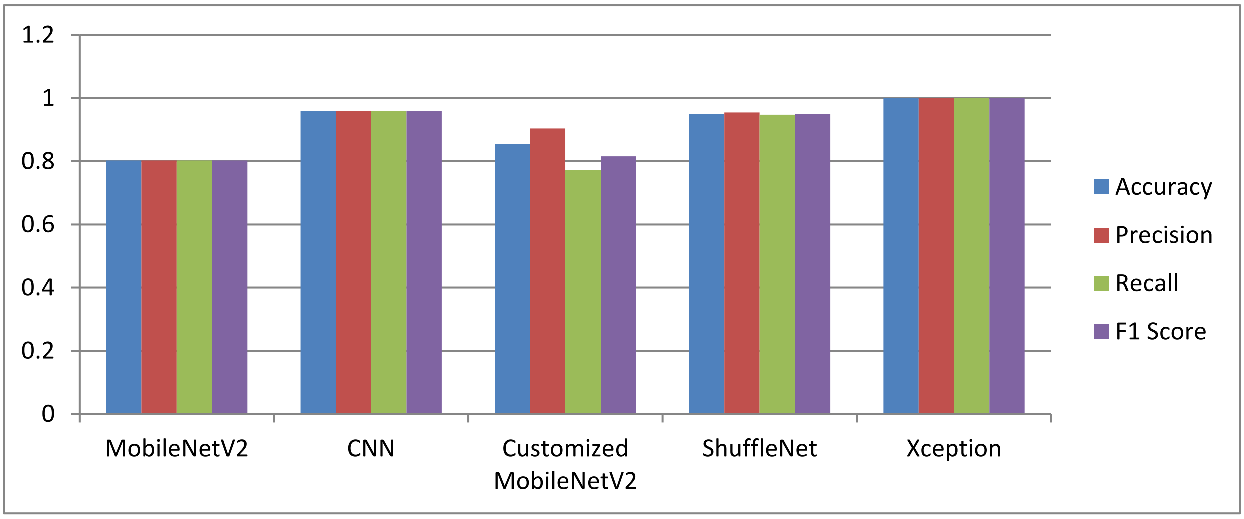
Table.5 Performance Evaluation Metrics of Classification of Young AffectNet HQ

Model	Accuracy	Precision	Recall	F1 Score
MobileNetV2	0.802	0.802	0.802	0.802
CNN	0.959	0.959	0.959	0.959
Customized MobileNetV2	0.855	0.904	0.772	0.816
ShuffleNet	0.949	0.954	0.947	0.949
Xception	1.000	1.000	1.000	1.000

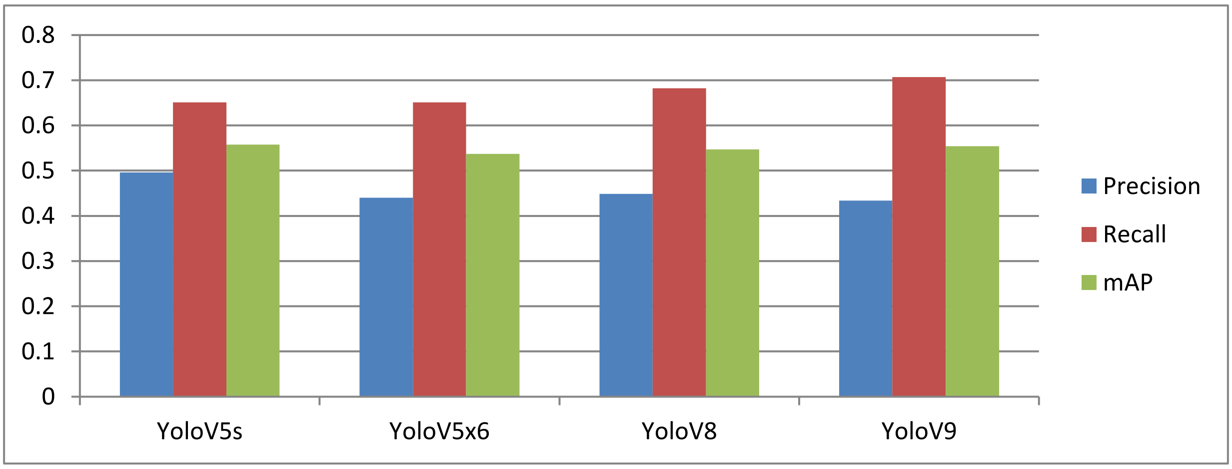
Table.6 Performance Evaluation Metrics of Detection of Young AffectNet HQ

Model	Precision	Recall	mAP
YoloV5s6	0.496	0.651	0.558
YoloV5x6	0.440	0.651	0.537
YoloV8	0.449	0.682	0.547
YoloV9	0.434	0.707	0.554

Graph 5 Comparison Graphs of Classification Young AffectNet HQ



Graph.6 Comparison Graphs of Detection Young AffectNet HQ



International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

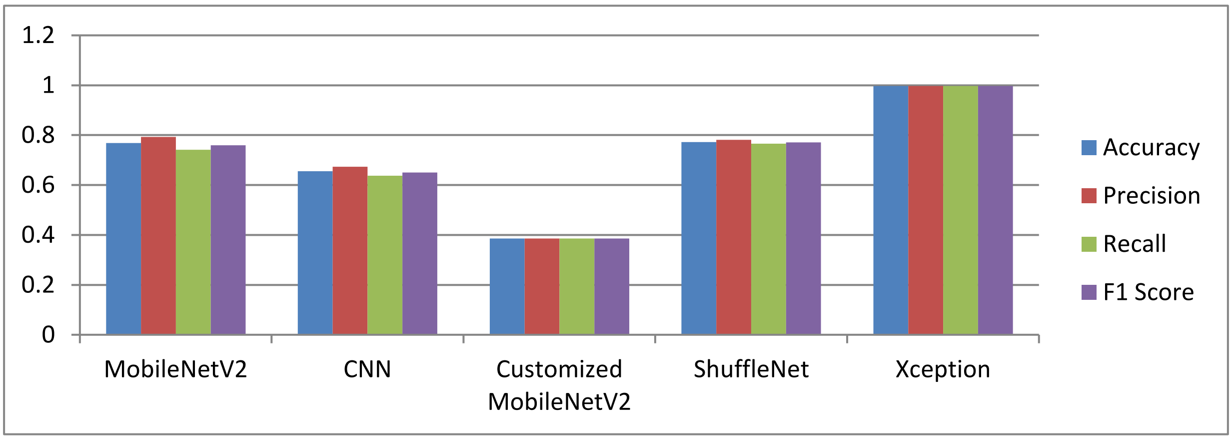
Table.7 Performance Evaluation Metrics of Classification of Young AffectNet HQ

Model	Accuracy	Precision	Recall	F1 Score
MobileNetV2	0.768	0.793	0.742	0.759
CNN	0.656	0.673	0.638	0.650
Customized MobileNetV2	0.386	0.386	0.386	0.386
ShuffleNet	0.772	0.781	0.766	0.771
Xception	0.997	0.997	0.997	0.997

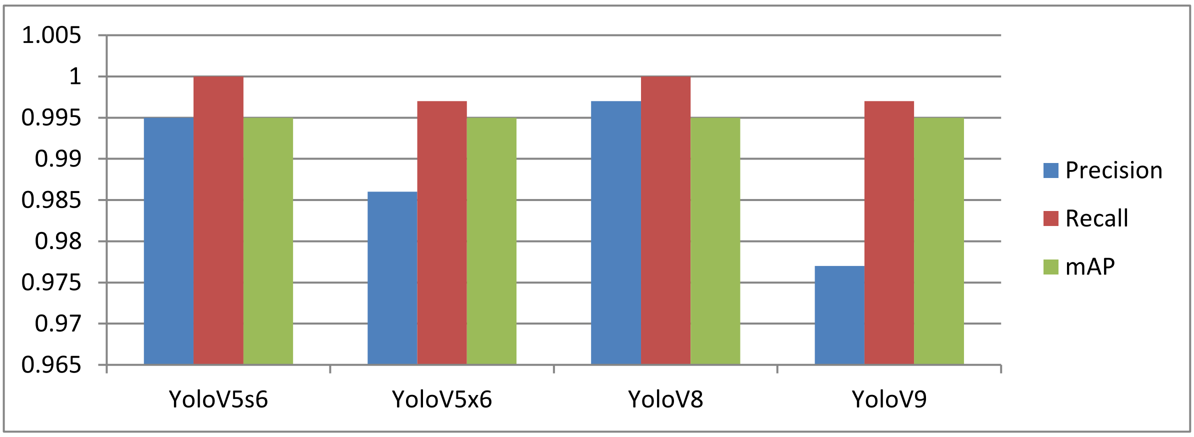
Table.8 Performance Evaluation Metrics of Detection of Young AffectNet HQ

Model	Precision	Recall	mAP
YoloV5s6	0.995	1.000	0.995
YoloV5x6	0.986	0.997	0.995
YoloV8	0.997	1.000	0.995
YoloV9	0.977	0.997	0.995

Graph 7 Comparison Graphs of Classification CK+ Dataset



Graph.8 Comparison Graphs of Detection CK+ Dataset



International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Accuracy (light blue), precision (maroon), recall (green), and F1-score (violet) are shown in classification graphs. Precision in detection graphs is represented by light blue, recall by maroon and mAP by green (Graphs 1 to 8). Xception and YOLO are more demonstrative in performance, as they provide the highest values in all cases compared to the other models. These findings are graphically illustrated in the graphs above.

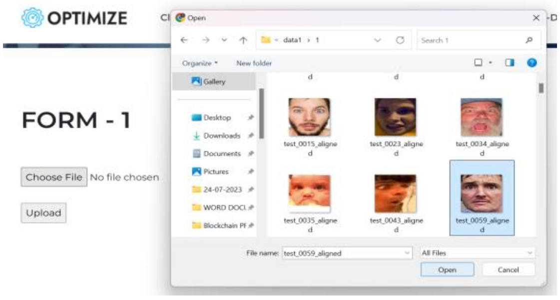


Fig. 6 Input Page

The facial emotion recognition application interface is presented in figure 6, which comprises a file selection dialog box that holds the pictures of faces.

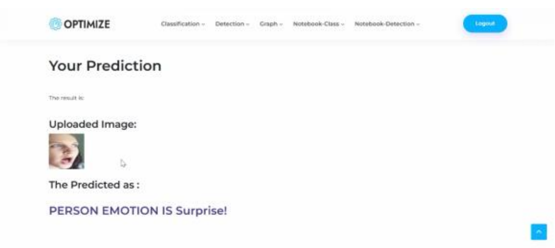


Fig. 7 Predicted Output

The result of a face emotion detector system is shown in Figure 7 which displays a posted photo and the expected emotion of surprise!.

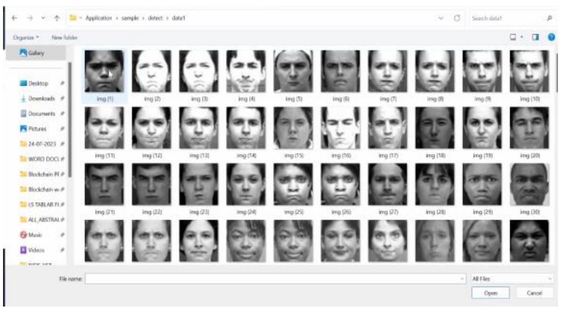


Fig. 8 Input page

Figure 8 shows a file selection dialog that has a grid of a few cropped grayscale faces probably used in a facial recognition or analysis task.

Fig. 9 Output page

The picture in figure 9 is a cropped image of a face with a green bounding box around it and the emotion is identified as anger with a confidence of 0.9.

5. CONCLUSION

This study on FER has effectively utilized advanced ML models, such as MobileNetV2, Customized MobileNetV2, ShuffleNet, and Xception, to get superior emotion detection performance. These are light and lean algorithms that have been optimized to achieve high precision and high-performance inference, making them suitable to be used in real-time. The fusion of YOLOv5x6, YOLOv8 and YOLOv9 greatly enhanced the ability of the system to detect and classify emotions on the faces of people in moving environments with amazing precision. An in-depth analysis of publicly available datasets, such as FER2013, RAF-DB, and CK-Dataset, showed that the proposed models are effective, and they significantly improve performance when compared to traditional methods. This shows the promise of ML techniques in many applications, including human-computer interaction and extending security

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

and surveillance systems, and underscores the necessity to make computing efficient in a real-world setting.

The next avenues of this study are to explore the incorporation of more advanced methods, including transfer learning and ensemble that can be used to enhance the performance of the FER systems by utilizing the pre-trained models and integrating numerous models to produce an improved precision. Further, the extension of the scope to include emotion detection based on textual and audio input along with multi-modal techniques can contribute to the usefulness and robustness of the system. Real-time feedback will facilitate the system to adapt to the evolving user needs, therefore, enhancing the level of engagement and accuracy. Further research on the impact of environmental factors, such as cultural differences in facial expression recognition will help develop a more flexible and universal emotion recognition system.

REFERENCES

[1] F. M. Talaat, Z. H. Ali, R. R. Mostafa, and N. El-Rashidy, ‘‘Real-time facial emotion recognition model based on kernel autoencoder and convolutional neural network for autism children,’’ Soft Comput., vol. 35, no. 17, pp. 1–14, Jan. 2024.

[2] M. Krumnikl and V. Maiwald, ‘‘Facial emotion recognition for mobile devices: A practical review,’’ IEEE Access, vol. 12, pp. 15735–15747, 2024.

[3] Z.-Y. Huang, C.-C. Chiang, J.-H. Chen, Y.-C. Chen, H.-L. Chung, Y.-P. Cai, and H.-C. Hsu, ‘‘A study on computer vision for facial emotion recognition,’’ Sci. Rep., vol. 13, no. 1, p. 8425, May 2023.

[4] K. Sarvakar, R. Senkamalavalli, S. Raghavendra, J. S. Kumar, R. Manjunath, and S. Jaiswal, ‘‘Facial emotion recognition using convolutional neural networks,’’ Mater. Today, Proc., vol. 80, pp. 3560–3564, Sep. 2023.

[5] E. Pranav, S. Kamal, C. S. Chandran, and M. H. Supriya, ‘‘Facial emotion recognition using deep convolutional neural network,’’ in Proc. 6th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS), Mar. 2020, pp. 317–320.

[6] S. Li and W. Deng, ‘‘Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition,’’ IEEE Trans. Image Process., vol. 28, no. 1, pp. 356–370, Jan. 2019.

[7] I. J. Goodfellow et al., ‘‘Challenges in representation learning: A report on three machine learning contests,’’ in Proc. Int. Conf. Neural Inf. Process., in Lecture Notes in Computer Science, M. Lee, A. Hirose, Z.-G. Hou, and R. M. Kil, Eds. Berlin, Germany: Springer, 2013, pp. 117–124.

[8] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, ‘‘The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression,’’ in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jun. 2010, pp. 94–101.

[9] L. K. McCorry, ‘‘Physiology of the autonomic nervous system,’’ Amer. J. Pharmaceutical Educ., vol. 71, no. 4, p. 78, Sep. 2007, doi: 10.5688/aj710478.

[10] J. S. Barlow, The Electroencephalogram: Its Patterns and Origins. Cambridge, MA, USA: MIT Press, 1993.

[11] M. A. Hasnul, N. A. A. Aziz, S. Alelyani, M. Mohana, and A. A. Aziz, ‘‘Electrocardiogram-based emotion recognition systems and their applications in healthcare—A review,’’ Sensors, vol. 21, no. 15, p. 5015, Jul. 2021. [Online]. Available: https://www.mdpi.com/1424- 8220/21/15/5015

[12] F. Ekman, ‘‘Facial action coding system,’’ in Environmental Psychology & Nonverbal Behavior. Palo Alto, CA, USA: Consulting Psychologists Press, 1978, doi: 10.1037/t27734-000.

[13] Y. Fan, J. C. K. Lam, and V. O. K. Li, ‘‘Demographic effects on facial emotion expression: An interdisciplinary investigation of the facial action units of happiness,’’ Sci. Rep., vol. 11, no. 1, p. 5214, Mar. 2021. [Online]. Available: https://www.nature.com/articles/s41598-021-84632-9

[14] R. Leung, ‘‘Subsurface boundary geometry modeling: Applying computational physics,

computer vision, and signal processing techniques to geoscience,'' IEEE Access, vol. 7, pp. 161680–161696, 2019.

[15] J. D. A. Dornelles, N. F. Ayala, and A. G. Frank, ''Smart working in Industry 4.0: How digital technologies enhance manufacturing workers' activities,'' Comput. Ind. Eng., vol. 163, Jan. 2022, Art. no. 107804. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835221007087

[16] Y. Cai, W. Zheng, T. Zhang, Q. Li, Z. Cui, and J. Ye, ''Video based emotion recognition using CNN and BRNN,'' in Proc. 7th Chin. Conf. Pattern Recognit., Chengdu, China. Tokyo, Japan: Springer, 2016, pp. 679–691.

[17] E. C. Nook, K. A. Lindquist, and J. Zaki, ''A new look at emotion perception: Concepts speed and shape facial emotion recognition,'' Emotion, vol. 15, no. 5, pp. 569–578, Oct. 2015.

[18] J. M. Leppänen and C. A. Nelson, ''The development and neural bases of facial emotion recognition,'' in Advances in Child Development and Behavior, vol. 34, R. V. Kail, Ed. San Diego CA, USA: JAI Publisher, Jan. 2006, pp. 207–246. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S006524070680008X

[19] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, ''Classifying facial actions,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 10, pp. 974–989, Oct. 1999.

[20] D. Mehta, M. F. H. Siddiqui, and A. Y. Javaid, ''Recognition of emotion intensities using machine learning algorithms: A comparative study,'' Sensors, vol. 19, no. 8, p. 1897, Apr. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/8/1897